
KOPAL

KOoperativer Aufbau eines Langzeitarchivs Digitaler Informationen



Ein Kooperationsprojekt

Agenda

1. Das KOPAL-Projekt: Eckdaten
2. Herausforderungen für ein Langzeitarchiv digitaler Informationen
3. OAIS Modell
4. Lösungskonzept DIAS
5. Lösungskonzept KOPAL
6. Projektziele
7. Projektpartner und deren spezifische Aufgaben / Sichten

Das KOPAL-Projekt

- Förderprojekt des BMBF
- Charakter: Forschungs- und Innovationsprojekt
- Fördervolumen: 4,2 Mio €, Laufzeit: 1.7.2004 – 30.6.2007
- Projektziel: Aufbau einer von Gedächtnisorganisationen nachnutzbaren technischen und organisatorischen Infrastruktur zur Sicherung der Langzeitverfügbarkeit elektronischer Publikationen
- Basis: DIAS (Digital Information Archiving System) entwickelt von IBM für die Koninklijke Bibliotheek, die Nationalbibliothek der Niederlande
- Projektpartner
 - Die Deutsche Bibliothek (DDB) (Gesamtprojektleitung)
 - Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB)
 - Gesellschaft für wissenschaftliche Datenverarbeitung (GWDG)
 - IBM Deutschland GmbH (IBM)

Herausforderungen eines Langzeitarchivs digitaler Informationen 1

- Auflösung heutiger Engpässe bei der Archivierung bzgl. Anzahl und Art der Dokumente
- Rasanter Technologiewechsel
 - erschwert Zugriff auf ältere Datenformate
 - erfordert Emulation der ursprünglichen Umgebung oder Migration der Daten
 - Zeitalter der Globalisierung fordert verbindliche Definitionen dazu, welche Informationen für die Langzeitarchivierung digitaler Ressourcen erforderlich sind und eines international einheitlichen Standards, zur Art und Weise der Verfügbarkeit
 - Forderung neuer Lösungen und Zwang zur Selektion durch die explosive Zunahme an Informationsmenge, versionierten Dokumenten, Annotationen an dynamischen Dokumentformaten
- Technische Aspekte: OAIS-Modell, Metadatenmodellierung, Authentifizierung

Herausforderungen für ein Langzeitarchiv digitaler Informationen 2

- Rasch wachsender elektronischer Publikationssektor
- Zunehmender Teil der kulturellen Aktivitäten im Netz
- Selektionskriterien entscheiden über die Tradierung unseren kulturellen Erbes
- Dieser „Nachlass“ ist – auch gesetzlich fixiert – Aufgabe von Nationalbibliotheken und vergleichbaren Einrichtungen
- Bestehende Archive erfüllen Anforderungen an „trusted digital repositories“ nicht

OAIS Modell

Funktionale Anforderungen, auf denen DIAS beruht:

- Die relevanten Business Prozesse basieren auf dem OAIS Modell (=Reference Model for an „Open Archival Information System“ - ISO 14721
- OAIS dient als Framework für die Lösung
- Der Kern von DIAS besteht aus 6 Prozessen, die mit EPs in den drei Standardformaten arbeiten:
 - Ingest Process
 - Archival Storage
 - Data Management
 - Access
 - Administration
 - Preservation

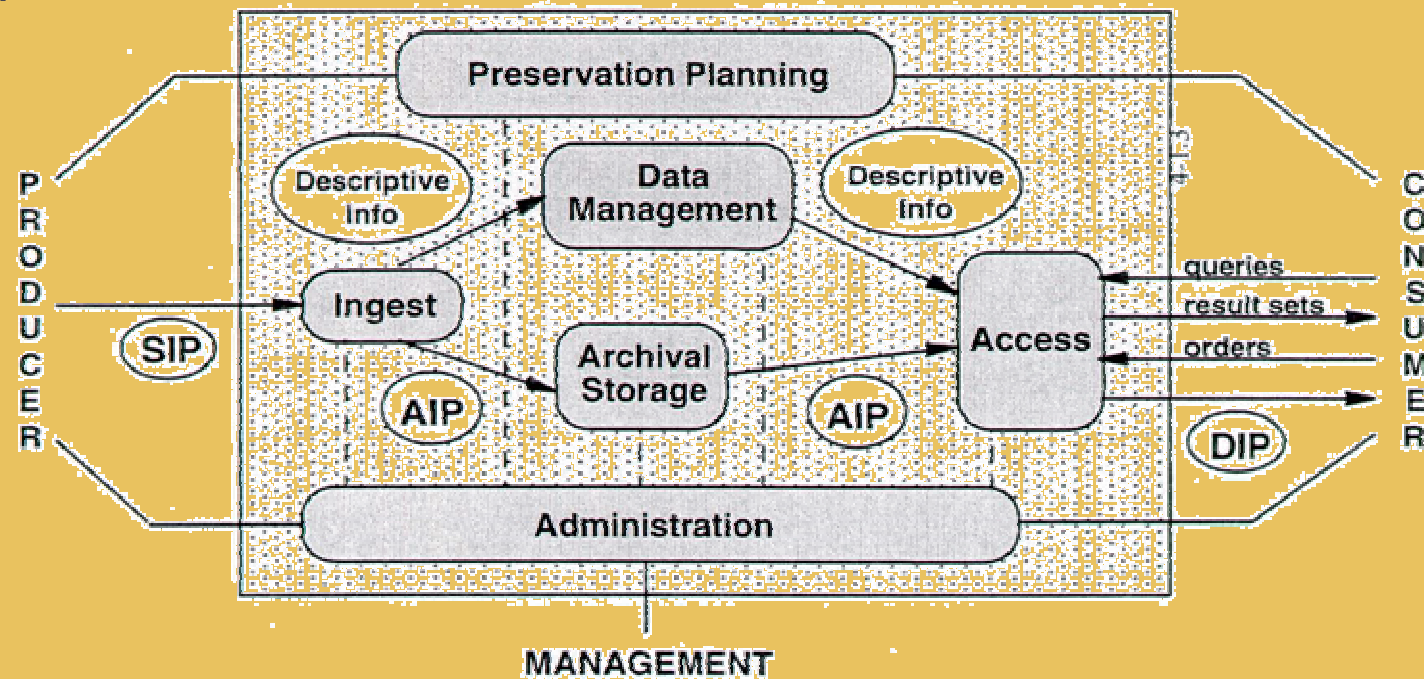
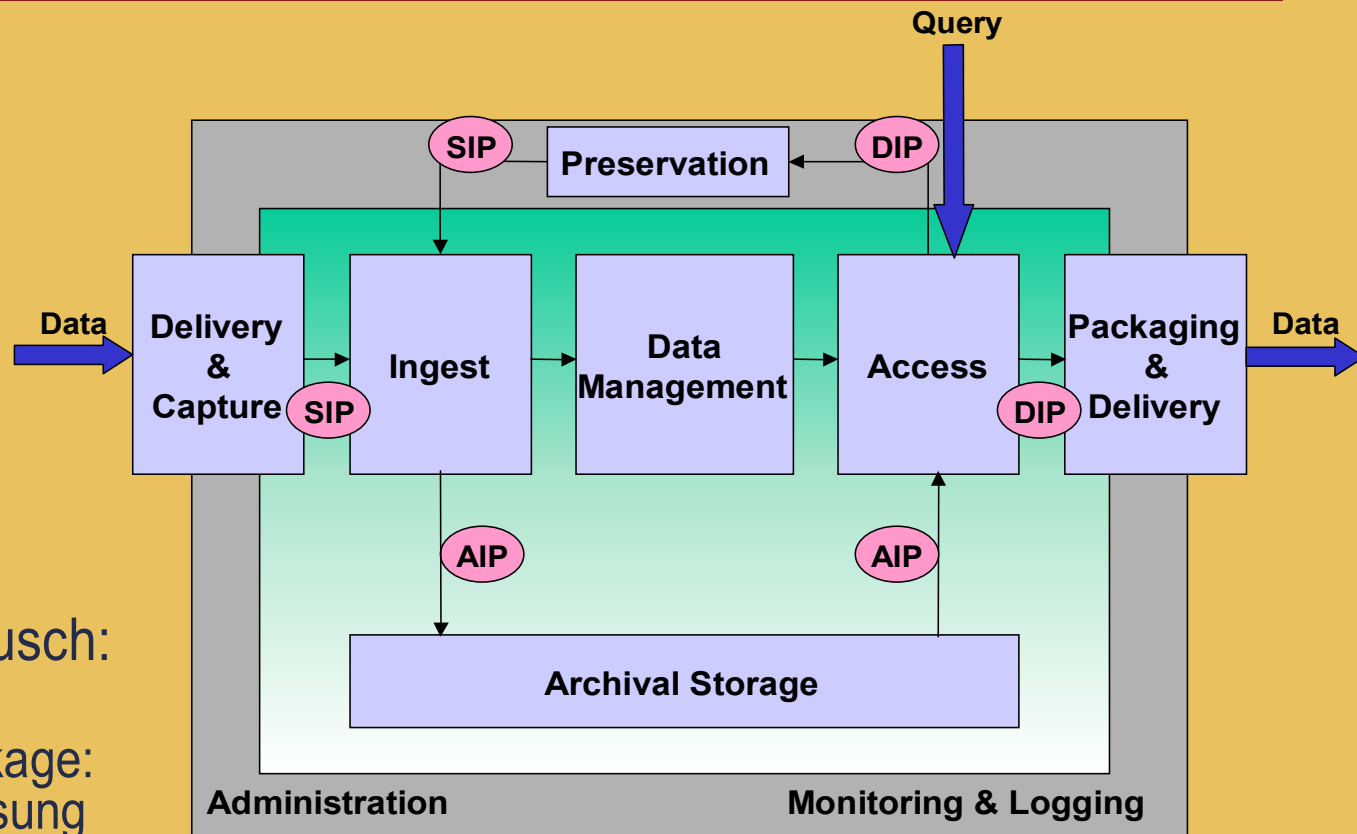


Figure 4-1: OAIS Functional Entities

- Basis ist das OAIS Modell
- Erweiterungen:
 - Delivery & Capture
 - Packaging & Delivery
 - Monitoring & Logging

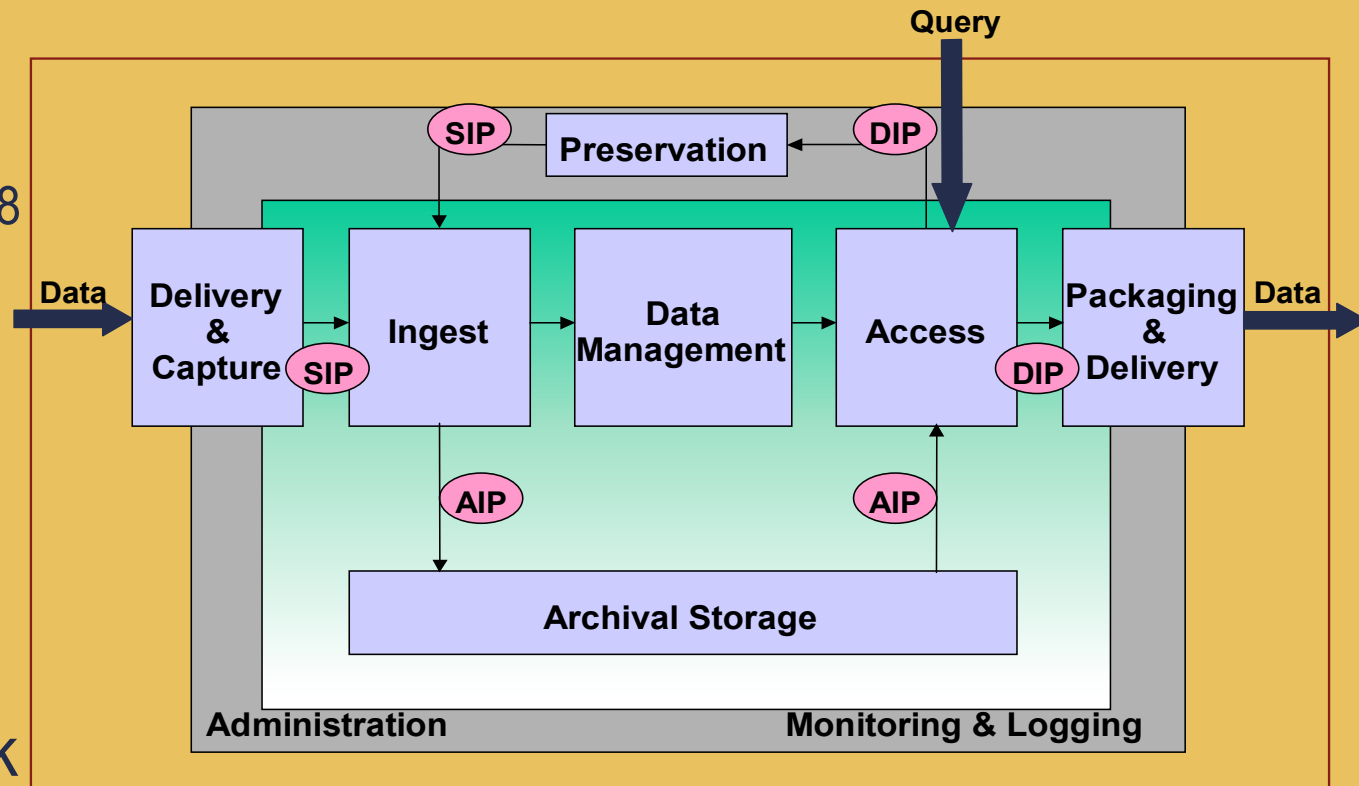
- 3 Standardformate zum Austausch:

- SIP = Submission Information Package: Container für die EP Einspeisung
- AIP = Archival Information Package: residente Darstellung der EPs (Metadaten + digitale Objekte)
- DIP = Dissemination Information Package: (Container für die Auslieferung von EPs an User kann neben Objekt auch Anw.SW enthalten)



Source: Networked European Deposit Library

- DIAS Erweiterung
 - Projektbeginn
 - Mandantenfähigkeit
 - Fernzugriff
 - DIAS Upgrade auf CM 8
 - Projektdurchführung
 - Funktionalitäten zu Preservation Manager
 - Funktionalitäten zu Preservation Planning
 - METS Support
- Open-Source Framework
 - Batch-Builder
 - Access Lösungen für graphische Darstellungen und Daten



Quelle: NEDLIB

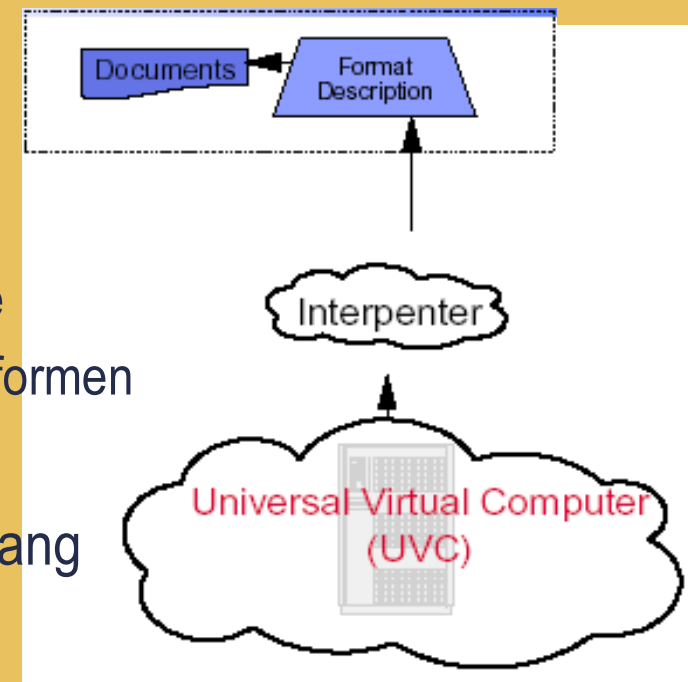
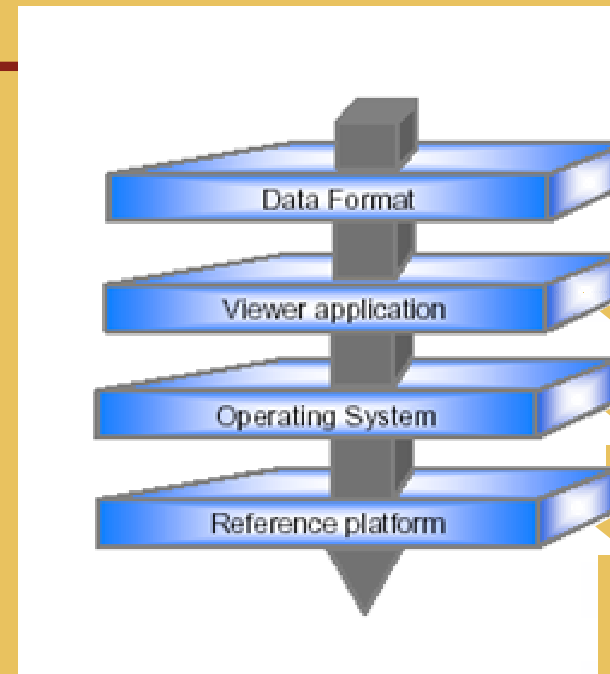
Preservation Planning

- Preservation Layer Model (PLM)

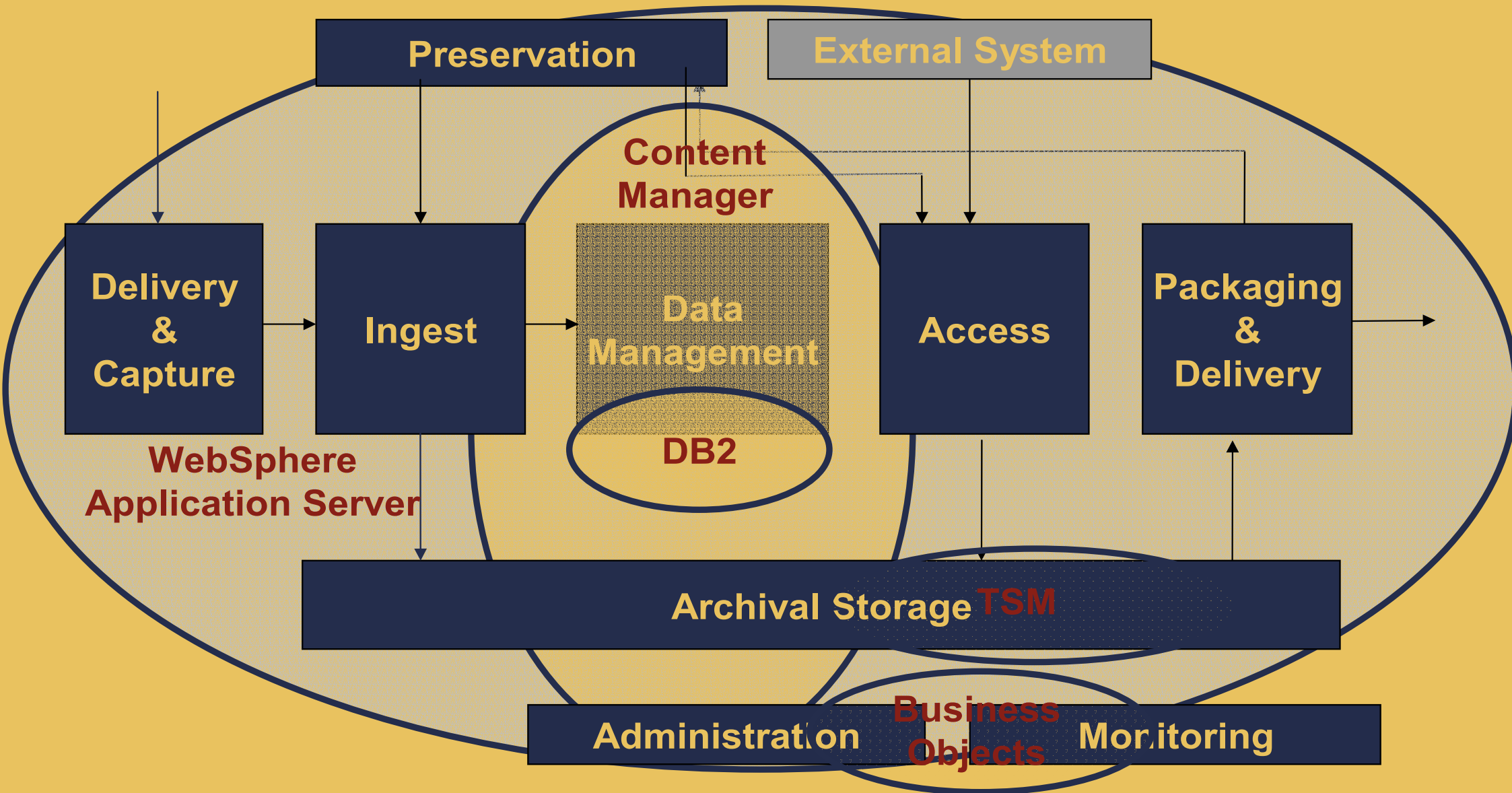
- Zugang unabhängig von technologischen Veränderungen
- basiert auf Beziehung zwischen Infrastrukturkomponenten (Zugang, Präsentation EP)
- Entwicklung von ViewPaths für jedes archivierte Datenformat

- Universal Virtual Computer (UVC)

- Emulationsansatz für flexiblen Emulator
 - Interpretation Semantik beider Datenformate
 - Emulation von Hardware-/Arbeitssystemplattformen
- Archivierung von EP in Standardformat
- Viewer mit UVC Technologie garantiert Zugang



Lösungskonzept von DIAS – Nutzung von Standard-Software



Ziele des KOPAL Projektes

- Langzeitarchivierung des kulturellen deutschen Erbes (>100Jahre)
- Entwicklung einer standardisierten Lösung
 - nachnutzbarer Transfer für andere Institutionen
 - vereinfachte Kooperation zwischen Bibliotheken
- Erweiterung des IBM Assets DIAS-Core um Open-Source basierte Komponenten
 - flexibler Einsatz durch Modularität der Lösung
- Entwicklung einer erweiterbaren User Group (KB, DDB, SUB) zum kooperativer Betrieb der Lösung
 - Niedrigere Kosten
 - Zeitlich schnellere Weiterentwicklung
 - Stärkere Standardisierung möglich
- Prozessintegration z.B. Automatisierung der Einlieferrouinen, Optimierung für Massenverfahren

Projektziel / Motivation

- Entwicklung einer tragfähigen Lösung für die Langzeitarchivierung der stark anwachsenden Zahl deutscher Netzpublikationen

Kompetenzen

- Zentrale Archivbibliothek Deutschlands
- Maßgeblich an der Entwicklung internationaler Standards beteiligt
- Umfangreiche Sammlung deutscher elektronischer Dissertationen und Netzpublikationen

Aufgaben

- Gesamtprojektleitung (R. Altenhöner)
- Konzeption technischer Anforderungen
- Einbringen der Erfahrungen bei Sammlung und Archivierung von Netzpublikationen
- Internationale Zusammenarbeit

Projektleitung DDB

- Tobias Steinke, steinke@dbf.ddb.de

Projektziel / Motivation

- Digitalisierte Wissens- und Informationsbestände im langfristigen Zugang für Wissenschaftler und Lernende bereitstellen

Kompetenzen

- Serviceorientierter Informationsdienstleister
- Innovationsorientierung durch mehr als 20 Projekte zur Digitalen Bibliothek
- GDZ als bundesweites Kompetenzzentrum für Digitalisierung und Metadatenhandling
- LRC für medienneutrales Publizieren

Aufgaben

- Erarbeitung der METS basierten Objektstruktur(en)
- Erstellung / Programmierung von Import- und Exportroutinen (TIFF, TeX)
- Entwicklung von Workflows
- Öffentlichkeitsarbeit

Projektleitung SUB

- Frank Klaproth, klaproth@sub.uni-goettingen.de

Projektziel / Motivation

- Langfristige Datenarchivierung ist eine Kernaufgabe der GWDG
- Langzeitarchivierung als künftig wichtige Dienstleistung der GWDG
- Großes Interesse an LZA auf Seiten beider Gesellschafter

Kompetenzen

- Rechen- und Kompetenzzentrum für Universität Göttingen und Max-Planck-Gesellschaft
- Backup- und Archivsystem mit räumlich getrennten Standorten der eingesetzten Bandbibliotheken
- Betreiber Hochgeschwindigkeitsnetz GÖNET (1 Gbit/sec) mit 622 Mbit/sec Anbindung an das Gigabit-Wissenschaftsnetz (G-WiN)

Aufgaben

- Hardware: Beschaffung + Pflege
- IT-Infrastruktur
- Serveradministration
- Speichermedienverwaltung

Projektleitung GWDG

- Dagmar Ullrich, dullric@gwdg.de

Projektziel / Motivation

- Einbringen von Expertise und Technologien für den Aufbau einer standard-basierten Langzeitarchivierungslösung für digitale Medien. Arbeiten mit innovativen Kunden, um die Lösung eng an den fachlichen Anforderungen auszurichten.

Kompetenzen

- spezifische Expertise durch Aufbau von DIAS bei der Nationalbibliothek der Niederlande
- Technologien für Langzeitarchivierung als Produktschwerpunkt (z.B. Universal Virtual Computer als Basis für Emulations- / Migrationsstrategien von elektronischen Publikationen)
- Service- und Projektkompetenz

Aufgaben

- Lieferung von DIAS-Core und IBM Content Manager als Basis zur Langzeitarchivierung
- Funktionale und technologische Weiterentwicklung von DIAS
- Organisation und Koordinierung der europäischen DIAS User Group

Projektleitung IBM

- Irmgard Kurth, ikurth@de.ibm.com

Back-up
